Journal of Electronic Research and Application, 2025, Volume 9, Issue 5
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on the Intelligent Evaluation of University Bursary Based on Blockchain

Juan Li*

Gannan University of Science and Technology, Ganzhou 341000, Jiangxi, China

**Author to whom correspondence should be addressed.*

Abstract: Aiming at the problems of easy falsification of information and inaccurate evaluation results in the existing university bursary evaluation, a bursary evaluation model (XGBoost Model based on Blockchain, XMB) combining machine learning and blockchain was designed. The relevant basic data of the bursary evaluation is stored on the chain to solve the problem of easy falsification of data in the evaluation process. At the same time, the evaluation results of the student bursary are uploaded to the chain to realize the traceability of historical data. In addition, the improved XGBoost algorithm is used to intelligently analyze and evaluate the basic data of students, and objectively give the student a bursary grade, which realizes the intelligence and scientific nature of the evaluation process and ensures the accuracy of the evaluation results. The experimental results prove that the model proposed in this paper has an accuracy rate of about 6% higher than that of the traditional XGBoost model, which has higher evaluation accuracy, throughput, and time efficiency. The method proposed in this paper is suitable for the evaluation of scholarships and bursaries in the student management system of colleges and universities.
Keywords: Blockchain; Bursary evaluation; Machine learning; XGBoost model

Online publication: October 21, 2025

1. Introduction

The management of university bursaries is an important task in the management of students in colleges and universities in our country, and it plays an important role in reflecting the country's care for students in financial difficulties and achieving the fair development of higher education. However, due to the continuous expansion of enrolment in colleges and universities across the country in recent years, the large number of students, and uneven regional economic development, there are some difficulties and problems in the management of university bursaries. For example, it is difficult to determine the qualifications of university bursaries, the assessment of university bursaries is inaccurate and unobjective, and the information on university bursaries has been tampered with.

In order to solve the above problems, researchers have applied machine learning technology to the management of university scholarships. Lu *et al.* [1] combined the XGBoost model and principal component analysis to establish a classification and prediction method for poor students in colleges and universities with high accuracy. Li [2] used the K-nearest neighbor algorithm to increase the accuracy of predicting the real poor students. However, the above scheme does not involve a data protection mechanism, and lacks data security protection and personal data privacy protection. At the same time, some researchers have proposed applying blockchain technology to the management of university scholarships and bursaries to solve the above-mentioned related problems. Blockchain is a technical architecture that integrates cryptography, game theory, computer science, and other disciplines. It includes specific technologies such as distributed storage, consensus algorithms, asymmetric keys, peer-to-peer networks, and smart contracts. The combination of these technologies enables the blockchain to have the characteristics of decentralization, a high degree of autonomy, safety, reliability, and non-tamperability, giving the blockchain unique functions that are different from traditional systems. In the management of university scholarships and bursaries, Ding *et al.* [3] built a blockchain scholarship management platform to record student-related evaluation data, and trigger smart contracts to generate evaluation results when the scholarship conditions are met; Yan [4] applied blockchain technology to create a bursary data sharing platform, which improved the credibility of data interaction and avoided the problem of data tampering and forgery in the bursary management process. However, the above scheme lacks data intelligent optimization processing and fails to solve problems such as inaccurate evaluation results.

In order to solve the above problems, this paper proposes a research plan for the intelligent evaluation model of university bursaries based on blockchain:

(1) Combined with the use of the machine learning XGBoost algorithm, intelligent model training, and poverty level result prediction on student-related data.

(2) Store student data and assessment results on the blockchain, and design related smart contracts to ensure that the data can be traced and cannot be tampered with, etc.

2. Intelligent evaluation model

Aiming at the difficulties in identifying the qualifications of university bursaries and the different evaluation reference standards, this paper will use the improved XGBoost model to predict and analyze the grade evaluation of poor students.

2.1. Introduction of XGBoost

XGBoost (eXtreme Gradient Boosting) is an integrated learning model based on the gradient boosting algorithm proposed by Chen *et al.* [5]. The model idea is to build an integrated model with strong learning ability by using a combination of multiple weak models. The model is continuously upgraded and updated. Each iteration continues to generate a new model on the basis of the generated model, so that the model is constantly approximating the sample distribution. XGBoost combines the first-order derivative with the second-order derivative, and the algorithm uses the tree model complexity as a regular term in the objective function to avoid overfitting [6].

The objective function of the XGBoost model is:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i \hat{y}^{(t-1)} + f_t(x_i) + \Omega(f_t) + C \quad (1)$$

Where $l(\bullet)$ is the loss function, $\Omega(f_t)$ is the regular term, used to control the complexity of the model and prevent over-fitting, and C is the inherent normal error. The regular term is defined as:

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \tag{2}$$

Among them, T is the number of leaf nodes, ω is the node weight, γ and λ are values between 0–1, which are used for the objective function to control the leaf nodes and the weight of each node.

Then use Taylor expansion to approximate the original goal. The second-order Taylor expansion is:

$$f(x+\Delta x) \simeq f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)\Delta x^2 \tag{3}$$

And define:

$$g(i) = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}^{(t-1)}) \tag{4}$$

$$h(i) = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}^{(t-1)}) \tag{5}$$

So as to sort out the substitution and get a new objective function:

$$Obj^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \tag{6}$$

Taking the partial derivative of , get:

$$\omega_j = -\frac{\sum_{i\in I,j} g_i}{\sum_{i\in I,j} h_i + \lambda} \tag{7}$$

Finally got:

$$Obj^{(t)} = -\frac{1}{2}\sum_{j=1}^{T}\frac{(\sum_{i\in I,j} g_i)^2}{\sum_{i\in I,j} h_i + \lambda} + \gamma T \tag{8}$$

Use the objective function to find the tree with the best structure. The smaller the value, the better the structure of the tree. Each time we try to add a split to an existing leaf node, we calculate the difference between the structure score before the branch and the structure score after the branch, which is called the gain calculation:

$$Gain = \frac{1}{2}\left[\frac{(\sum_{i\in I,L} g_i)^2}{\sum_{i\in I,L} h_i + \lambda} + \frac{(\sum_{i\in I,R} g_i)^2}{\sum_{i\in I,R} h_i + \lambda} - \frac{(\sum_{i\in I,j} g_i)^2}{\sum_{i\in I,j} h_i + \lambda}\right] - \gamma \tag{9}$$

2.2. XGBoost model optimization

The XGBoost algorithm can improve and optimize the model through objective function optimization and parameter adjustment.

This paper improves on the original XGBoost model, uses the third-order approximation of the loss function to improve the accuracy of the model, and performs a third-order Taylor expansion of the objective function, assuming:

$$p(i) = \partial^3_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \tag{10}$$

Take the derivative of equal to 0 and get a new result:

$$\omega_j = -\frac{\sum_{i \in I,j} h_i + \lambda \pm \sqrt{\left(\sum_{i \in I,j} h_i + \lambda - 2\sum_{i \in I,j} g_i \sum_{i \in I,j} p_i \right)}}{\sum_{i \in I,j} p_i} \tag{11}$$

And suppose:

$$\Delta = \sqrt{\left(\sum_{i \in I,j} h_i + \lambda - 2\sum_{i \in I,j} g_i \sum_{i \in I,j} p_i \right)} \tag{12}$$

And then get the new objective function:

$$Obj^{(t)} \approx \frac{\left(-\sum_{j=1}^{T} g_i (h_i + \lambda) \pm \Delta \right) p_i - \frac{1}{2}\left(\sum_{j=1}^{T} (h_i + \lambda) \pm \Delta \right)^2 - \frac{1}{6}\left(\sum_{j=1}^{T} (h_i + \lambda) \pm \Delta \right)^3}{p_i^2} + \gamma T \tag{13}$$

After adopting the third-order Taylor expansion, the complexity of the model is increased, and the prediction and classification ability of the model is also enhanced.

2.3. XGBoost modeling

The XGBoost model establishment process is divided into six stages: reading data, setting parameters, training model, predicting results, saving model, and model calling. Among them, parameter setting and optimization are the key steps.

(1) Reading data: XGBoost can load training data in a variety of data formats, such as text data in libsvm format, two-dimensional numpy arrays, and XGBoost binary cache files. The data selected in this paper are from a college student's campus card consumption data, family economic surveys, etc. Before loading the data, it needs to be cleaned, summarized, feature selection, data discretization, missing values and outlier processing, etc. In addition, it is necessary to fully consider the data indicators (characteristics) that can reflect poor students in these data. Here, we select the total number of personal consumption, total consumption amount, average consumption amount per time, average daily consumption amount, etc.

(2) Setting parameters: The parameter settings are shown in **Table 1**. This is a gradual process and requires constant debugging to find the optimal tree structure.

(3) Training model: After parameter setting and optimization are completed, the core of modeling is model training. Read the relevant data of the students, divide the training set and the test set according to the ratio of 7:3, and use the train function class to train 70% of the data to obtain a poor student grade prediction model.

(4) Predicting results: After the model is trained, you can use the trained model to predict the test data, that is, use the predict function to predict.

(5) Saving model: Save the model by using the pickle function of the library to save and call the model in Python, and complete the model on-chain operation in the smart contract layer.

(6) Model calling: After the model is saved and loaded on the chain, the data can be reimported for predictive analysis when in use.

Table 1. Parameter list

Parameter settings	Specific description
N_estimators=200	The number of weak classifiers in the ensemble algorithm, and the experiment is set to 200
Eta=0.1	The learning rate in the integration, the default is 0.3, the value range is [0,1]
Max_depth=6	The maximum tree depth of the weak classifier, the default is 6
Objective =multi:softmax	Specify the learning objective function and learning task, and actually solve the minimum value of the generalization error
Gamma=20	The descent of the objective function required for further branching on the leaf nodes of the tree, range [0,+∞]
booster=gbtree	It stands for weak evaluator, gbtree stands for gradient boosting tree
Subsample=1	The proportion of the sample drawn during random sampling, the range is (0,1], the default is 1

3. Blockchain-based bursary evaluation model

This article built a university bursary evaluation system based on Hyperledger Fabric [7] to meet the decentralized management of university bursary management and improve management efficiency, and to ensure that student user-related information is not leaked or tampered with.

3.1. Model architecture design

The blockchain-based bursary assessment model design is shown in **Figure 1**, which is divided into data layer, smart contract layer, and service layer. In the data layer, the blockchain uses HDFS (Hadoop distributed file system) [8]. As the underlying platform for storing massive amounts of data, HDFS can store massive amounts of structured and unstructured data, and is suitable for university student bursary data application scenarios. In the smart contract layer, functions such as chaining and calling of the XGBoost model are implemented. When the service layer sends a contract call request to the smart contract layer, the smart contract layer verifies the permissions and interacts with the data layer. After completing the data-related operations, the smart contract layer will return the processing result.

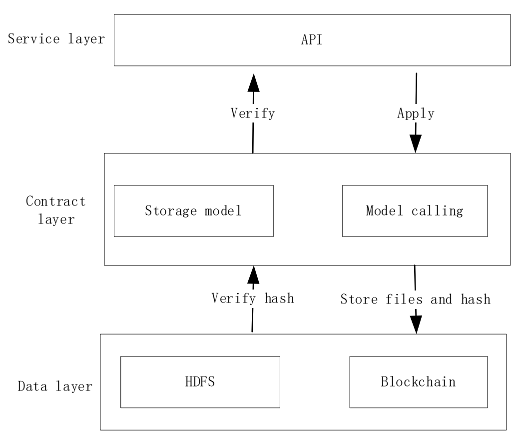

Figure 1. Blockchain-based bursary evaluation model

3.2. Smart contract design

The essence of a smart contract is that a piece of code is event-driven, using the agreement and user interface to complete the automatic execution of the contract. After the agreement is formulated and deployed, it can realize self-execution and verification without any peripherals or human intervention. In the Hyperledger Fabric platform, smart contracts are called chaincodes, which refer to application codes written in programming languages that provide state processing logic for distributed ledgers. In the smart contract layer of the system, two main functions are designed: XGBoost model store to blockchain and XGBoost model calling.

The model-on-chain contract deploys the bursary evaluation model that has achieved good results in training and testing on the blockchain. The smart contract on the chain of the model is as follows:

```
Algorithm 1: Deploy the model to the blockchain
Input: 1)ModelInfomation;2)TrainData;3)TestData
Output: 1)ResponseResult
    1.InputData:PackageModel(Parameters,TrainData,TestData);
    2.ArgsValidation:CheckArgs(InputData);
    3.Function ChaincodeInvoke(Operation,InputData);
    4.  if Operation=Storage;
    5.  if (Verify(Args) ==True;
    6.          StorageModel();
    7.          Return Success;
    8.  else;
    9.          Return Error;
    10.         end if
    11.  end if
```

After the XGBoost model is successfully connected to the chain, the service layer can call the model by calling the model contract. The model calls the smart contract as follows:

```
Algorithm 2: Invoke the model from the blockchain
Input: 1)UserId;2)Password;3)Function(Args);4)StudentData
Output: 1) valuationResult
    1. InputData: UserId, Password, Fun(Args), StundentData
    2. InformationValidation:CheckArgs(InputData);
    3. Function ChaincodeInvoke(Operation,InputData);
    4. if Operation=ModelInvoke;
    5.  if Model==Null;
    6.  Return True;
    7.  else;
    8.  Return ModelInvokeResult;
    9.  Return EvaluationResult;
    10. end if
    11. end if
```

4. Analysis of experimental results

The simulation experiment built a blockchain-based intelligent evaluation system for university scholarships. The experimental environment was Hyperledger Fabric version 2.1.1, Docker 19.2.1 version, Java 1.7 version, and Hadoop CDH3. Four hosts were configured, and 100 Docker nodes were set up. The experimental machine was configured as a CentOS 7.1 system, an Intel i7-5700k processor, 16 GB of memory, and JMeter 3.3 was used for stress testing.

4.1. Improved XGBoost model comparison

The basic XGBoost algorithm uses the second-order approximation method of the loss function to solve the objective function. This paper uses the basic XGBoost algorithm to solve the Taylor third-order expansion. As shown in **Figure 2**, the experiment shows that the accuracy rate [9] of the basic XGBoost model is between 70% and 80%, while in the improved XGBoost model, the accuracy rate reaches more than 80% in different test data scales. With the increase in test data, the accuracy rate of the methods shows an overall upward trend. The reason is that the improved XGBoost algorithm is based on the Taylor second-order expansion of the basic XGBoost algorithm by performing a third-order expansion to approximate the target item. Through more iterative updates, its accuracy is improved, and a better model effect is achieved.

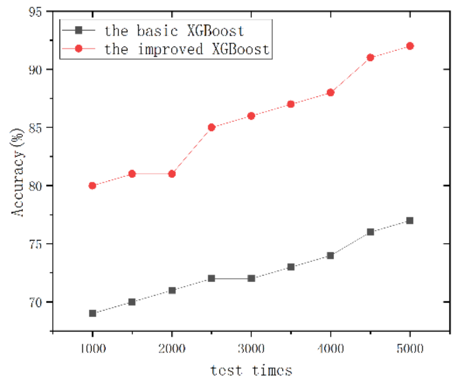

Figure 2. Comparison of XGBoost

4.2. Performance analysis

Insufficient performance is one of the challenges facing the blockchain. Throughput transactions per second (TPS) and response time are two key performance indicators that measure the blockchain system. Throughput refers to the number of transactions completed within a fixed time, and response time refers to the processing time to complete the transaction. In the TPS test, the original fabric model was compared with the model in this paper (XMB), and 50 and 100 nodes were set up for analysis and comparison, as shown in **Figures 3** and **4**. From the experimental data, when the number of nodes is 50, the number of XMB transactions per second reaches more than 500; when the number of nodes is 100, the number of XMB transactions per second reaches more than 700. No matter whether the node is 50 or 100, the TPS of XMB is higher than that of the native fabric system. The reason is that the model in this paper uses HDFS to store data in the data layer. Its advantage is that it supports massive data storage and has a large throughput; the state database that the native Fabric system can support is LevelDB and CouchDB, which use key-value pairs as storage methods. These two databases support complex queries, but their throughput performance is not as good as HDFS.

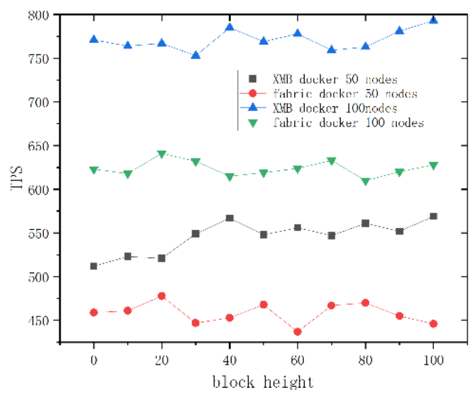

Figure 3. Comparison of TPS

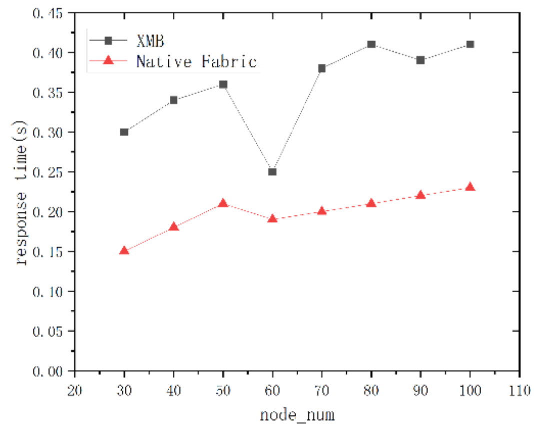

Figure 4. Response time comparison

In terms of response time, XMB has a slightly longer response time than native fabrics, because HDFS has the disadvantage of not being able to access data with lower latency [10]. But overall, XMB improves transaction processing efficiency and has less impact on time performance.

5. Conclusion

Aiming at the problems of traditional bursary evaluation, such as difficult qualification and inaccurate evaluation of university bursaries, this paper proposed a bursary evaluation model based on blockchain and machine learning algorithms, which realizes intelligent and scientific bursary evaluation. The non-tamperable and traceable basic data explores the innovative application of blockchain technology in the teaching field. Through in-depth analysis, it can be found that blockchain and artificial intelligence have a natural intergrowth and symbiosis [11]. Artificial intelligence provides optimization strategies for the core technology of blockchain, and blockchain can provide trust mechanism guarantees and infrastructure for artificial intelligence. In the future scenarios of blockchain digital infrastructure construction, most of which are faced with vertical structure scenarios such as government affairs, medical care, Internet of Things, education, etc., which need to be integrated into artificial intelligence, big data, and other fields, so artificial intelligence and blockchain are studied. The cross-integration of the company has a wide range of application prospects. The next step is to consider applying the model to other areas of university teaching and improving the data processing and optimization capabilities of the model.

Disclosure statement

The author declares no conflict of interest.

References

[1] Lu GM, Zhang Y, Zhou ZM, 2019, Research on Classification and Prediction of Poor Students Based on Machine Learning. Computer Applications and Software, 36(1): 316–319.

[2] Li MJ, 2017, Research on the Identification Method of Poverty Grants Based on Data Mining, Master's thesis,

Central China Normal University.

[3] Ding ML, Pang ST, Shang GY et al., 2020, A Blockchain-Based Method, Equipment and Medium for Evaluation and Scholarship Management, Innovation: CN111310200A.

[4] Yan XY, 2020, A Financial Aid Management Method and System Based on Blockchain Technology, Innovation: CN111797162A.

[5] Chen TQ, Guestrin C, 2016, XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[6] Torlay L, Perrone-Bertolotti M, Thomas E, et al., 2017, Machine Learning-XGBoost Analysis of Language Networks to Classify Patients with Epilepsy. Brain Informatics, 4(3): 159–169.

[7] Androulaki E, Barger A, Bortnikov V, et al., 2018, Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains, Proceedings of the 13th EuroSys Conference, 1–15.

[8] Bao GB, Yu CJ, Zhao H, 2013, Researching on the Placement of Data Replicas in the System of HDFS Cloud Storage Cluster, China Intelligent Automation Conference Proceedings 255, 259–269.

[9] Zhou JD, Yang Y, Zhang MJ, et al., 2016, Constructing ECOC Based on Confusion Matrix for Multiclass Learning Problems. Science China Information Sciences, 59(1): 135–148.

[10] Wu XY, Wang SH, Zhang YD, 2017, Summary of the Theory and Application of K-Nearest Neighbor Algorithm. Computer Engineering and Applications, 53(21): 1–7.

[11] Liu Y, Yu FR, Li X, 2020, Blockchain and Machine Learning for Communications and Networking Systems. IEEE Communications Surveys & Tutorials, 22(2): 1392–1431.